

일반논문-06-11-1-07

동공과 글린트의 특징점 관계를 이용한 시선 추적 시스템

박진우^{a)b)}, 권용무^{a)†}, 손광훈^{b)}

Gaze Tracking System Using Feature Points of Pupil and Glints Center

Jinwoo Park^{a)b)}, Yong-Moo Kwon^{a)†}, and Kwanghoon Sohn^{b)}

요 약

본 논문에서는 한 대의 카메라와 푸르킨에 영상을 이용한 간편한 2차원 시선 추적 시스템을 제안한다. 이 시스템은 사용자의 한쪽 눈 영상을 얻기 위해 적외선 필터가 장착된 카메라와, 사용자가 모니터 상에 바라보고 있는 응시 점을 알아내기 위해 각막의 표면에 반사점을 만들기 위한 두 개의 적외선 광원이 사용되었다. 카메라나 적외선 광원, 사용자의 머리는 자유롭게 움직일 수 있다. 따라서 본 시스템은 여타 불편한 고정된 장치나 사용자의 머리 고정이 필요 없는 간단하고 유연성 있는 시스템이다. 본 시스템은 또한 간편하고 정확한 사용자 캘리브레이션 과정을 포함하고 있다. 시스템을 사용하기에 앞서, 각 사용자는 시스템이 시선 추적 알고리즘 상의 개인 요소들을 초기화할 수 있도록 두 개의 점을 잠시 바라보기만 하면 된다. 제안된 시스템은 XGA (1024x768) 해상도에서 10 fps 이상 실시간으로 동작된다. 3명의 피 실험자와 9개의 실험 물체로 진행된 실험 결과는 시스템이 평균 1도의 시선 추적 오차를 보여 주고 있다.

Abstract

A simple 2D gaze tracking method using single camera and Purkinje image is proposed. This method employs single camera with infrared filter to capture one eye and two infrared light sources to make reflection points for estimating corresponding gaze point on the screen from user's eyes. Single camera, infrared light sources and user's head can be slightly moved. Thus, it renders simple and flexible system without using any inconvenient fixed equipments or assuming fixed head. The system also includes a simple and accurate personal calibration procedure. Before using the system, each user only has to stare at two target points for a few seconds so that the system can initiate user's individual factors of estimating algorithm. The proposed system has been developed to work in real-time providing over 10 frames per second with XGA (1024x768) resolution. The test results of nine objects of three subjects show that the system is achieving an average estimation error less than 1 degree.

Keyword: Gaze Tracking, Gaze Estimation, Pupil and Glint Detection, Human-Computer Interaction, Gaze-Based Interaction

I. 서 론

- a) 한국과학기술연구원 영상미디어연구센터
Imaging Media Research Center, Korea Institute of Science and Technology
- b) 연세대학교 전기전자공학부 디지털 영상미디어 연구실
Digital Image Media Lab, Department of Electronic and Electrical Engineering, Yonsei University
- † 교신저자 : 권용무(ymk@kist.re.kr)

미래의 유비쿼터스 환경을 위한 인간-컴퓨터 상호작용의 중요성이 높아짐에 따라 수많은 인간-컴퓨터 인터페이스가 사람과 컴퓨터를 자연스럽게 연결하기 위해 연구되고 개발되어지고 있다. 이러한 인터페이스 중에서도

컴퓨터 비전 기술을 이용한 것들은 카메라 외에는 어떤 다른 특별한 장치도 필요치 않다는 사실은 큰 장점으로 작용한다.

그리고 인간을 위한 인터페이스로써 시선은 인간의 관심이나 즉각적 반응, 흥미도, 그리고 욕구 등을 표현할 수 있는 효과적인 도구로써 이용될 수 있다.

본 논문에서 우리는 인간-컴퓨터 상호작용을 위한 입력 장치로써 시선을 사용하는 방법에 대해 연구하였다. 한 사람이 스크린 상의 점을 주시하고 싶을 때, 그는 그 점의 영상이 망막의 특정 부분인 중심와에 위치하도록 안구를 조절한다. 이 중심와는 사람이 어떤 물체에 초점을 맞출 때 시신경이 밀집해 있어 가장 선명한 상을 볼 수 있게 해주는 부분이다. 시선 추적이란 이런 스크린 상의 응시하고 있는 점을 눈 영상을 가지고 판별해 내는 것을 말한다.

현존하는 시선 추적 기술들은 2차원 추적 기술, 모델 기반 3차원 추적 기술, 그리고 3차원 추적 기술과 같이 세 개의 방법으로 구분되어질 수 있다^[1].

첫째로 2차원 추적 기술은 눈의 3차원 위치를 모르는 상태에서 사용자의 머리 움직임에 따른 눈의 상대적 방향만을 측정한다. 따라서 2차원 기술은 사용자의 머리를 고정시킬 필요가 있다. 이러한 방법은 대표적으로 전극을 눈 주위에 부착해 그 전위차를 이용한 EOG 방식^[2], 카메라 광축 주위에 부착된 광원과 광축에서 떨어진 광원을 교차로 사용하여, 이 때 동공이 밝고 어둡게 나오는 현상을 이용해 동공을 추적하는 방식^[3], 동공과 각막에 반사된 반사점을 이용해 시선을 추적하는 방식^[4], 1차와 4차 푸르킨에 영상을 이용해 시선 추적을 하는 듀얼 푸르킨에 방식^[5], 특수한 콘택트렌즈를 사용해 시선을 추적하는 서치 코일 방식^[6] 등이 있다.

두 번째로 모델 기반 3차원 추적 기술은 영상에서 3차원 시선 방향의 정보를 알아내기 위해 얼굴의 특징 점들을 추출하여 3차원 얼굴 모델을 기반으로 시선을 추적한다. 이 방법은 순수 컴퓨터 비전만을 이용한 방식으로 Seeingsmachine사에서 상용화시킨 "faceLab"이 있다^[7].

마지막으로 3차원 추적 기술은 헤드 트래커나 스테레오 카메라를 이용하여 눈의 3차원 위치를 추적한 후 2차원 추

적 기술을 적용하여 시선을 추적 한다^[1]. 팬-틸트-줌 카메라 등을 이용하여 카메라의 시야각에 의한 제한을 완화시키는 경우도 있다^{[8],[9]}.

본 논문에서는 위에 열거된 기술 중 동공과 각막의 반사점을 이용하는 방식으로써 한 대의 카메라와 두 개의 적외선 광원 집합을 사용한 간단한 2차원 시선 추적 기술을 제안한다. 대부분의 시선 추적 기술들이 3차원 정보를 이용하는 것과 달리 이 시스템은 단순히 눈 영상으로부터의 2차원 정보만을 사용하기 때문에, 본 시스템에서는 푸르킨에 영상을 얻을 수 있는 한도 내에서 적외선 광원이나 카메라, 사용자의 머리를 움직일 수 있다. 또한 시스템 구성이 간단하여 쉽게 구성이 가능하고, 구동 전의 복잡한 카메라 캘리브레이션 과정이 필요 없다. 사용자 캘리브레이션 방법 또한 매우 간편하여, 사용자는 본 시스템을 사용하기 전에 스크린의 좌 상단과 우 하단에 차례로 나타나는 두 점을 순서대로 응시하면 된다. 이러한 여러 가지 시스템의 유연성을 유지하면서도 시스템 오차는 사용하기에 충분히 높은 1도 이하의 정확도를 보여준다.

II. 시스템 개략

본 시스템은 한 대의 카메라와 두 개의 적외선 광원 집합으로 구성되어 있다. 두 광원은 사용자의 각막 표면에 푸르킨에 영상이라 불리는 두 개의 반사점을 생성한다. 시스템은 두 반사점과 동공의 중심을 이용하여 시선을 추적한다.

시선 추적 시스템 절차는 간단히 다음과 같은 4 단계로 나눌 수 있다. 시스템은 먼저 IEEE 1394 카메라를 사용하여 사용자의 눈 영상을 얻는다. 다음으로 획득한 눈 영상에서 데시메이션, 필터링, 프로젝션 히스토그램, 경계 검출 같은 영상 처리 기법을 사용하여 동공과 반사점의 중심들을 추출한다. 이렇게 구한 동공의 중심과 두 반사점의 중심을 세 특징 점으로, 이들 간의 거리 정보를 이용하여 사용자의 스크린 상 시선 위치로 매핑 시키게 되고, 이 점이 사용자 주시점으로 판별되게 된다. 이 모든 과정은 다음의 그림 1에 도시되어 있다.



그림 1. 시선 추적 절차
Fig. 1. Gaze Tracking Procedure

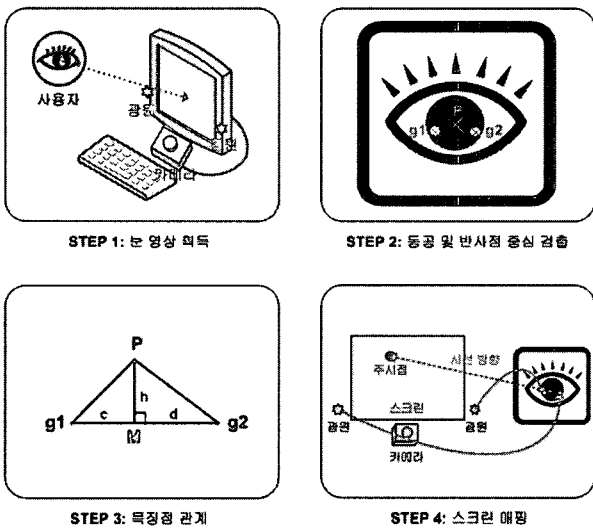


그림 2. 시선 추적 절차
Fig. 2. Gaze Tracking Procedure

III. 전처리 과정

스크린 상 주시점을 예측하기 위한 첫 번째 과정으로 사용자의 눈 영상이 얻어지면 해당 영상에서 동공의 중심과 두 반사점 중심을 검출해 내야 한다. 동공의 중심과 두 반

사점의 중심을 획득하는 전체 과정은 그림 3에 흐름도로 설명되어 있다. 먼저 시스템은 사용자의 선명한 눈 영상을 카메라를 통해 얻어내게 된다. 이 영상에서 필요한 정보는 단지 동공과 두 반사점이다. 눈 영상은 먼저 이진화 되고,

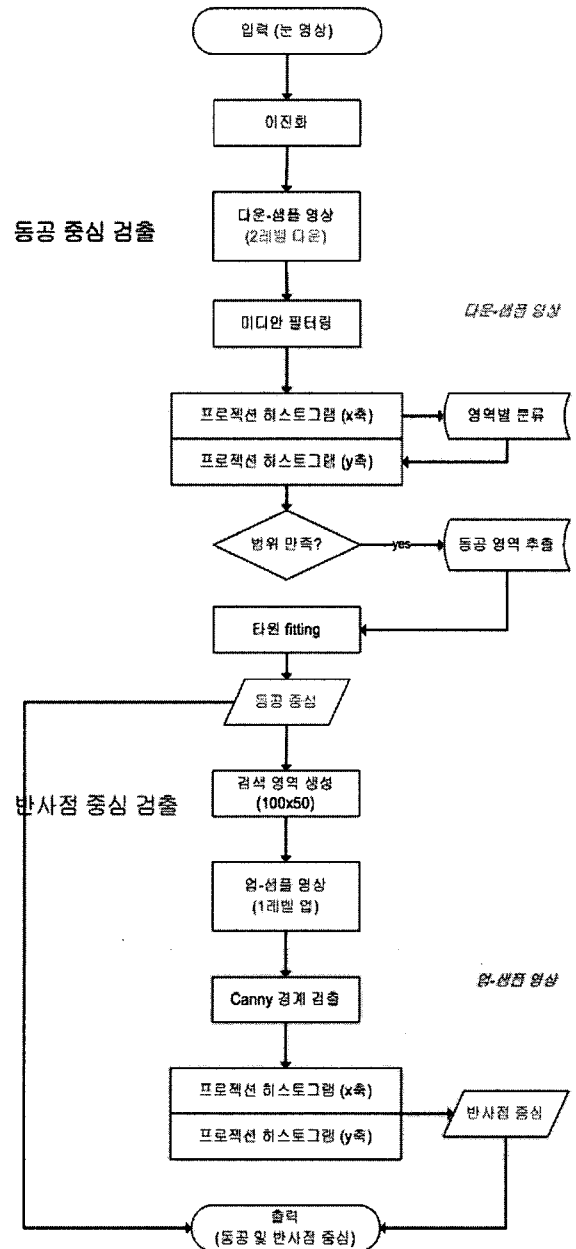


그림 3. 동공 및 반사점 중심 검출과정
Fig. 3. Detecting Pupil and Glints Center

가우시안 피라미드 분해를 통해 다운-샘플 되어 일종의 저대역 통과 필터로써 불필요한 노이즈가 제거된다. 이 과정에서도 남아 있는 노이즈는 미디안 필터를 이용해 추가로 제거하고, 해당되는 동공의 영역을 프로젝션 히스토그램을 사용해 추출하여 그 중심을 판별하게 된다. 동공의 중심을 알아내면 그 주위의 일정 부분에 반사점 중심을 알아내기 위한 검색 영역을 형성한다. 검색 영역 내의 영상은 다시 가우시안 피라미드 분해를 통해 업-샘플 되어 Canny 경계 검출을 통해 경계 영역만을 확보한다. 마지막으로 프로젝션 히스토그램을 사용하여 동공의 중심점 두 개를 판별하고, 찾아낸 동공의 중심과 반사점의 중심은 이후의 시선 추적 알고리즘의 입력으로 주시점을 판별하는데 이용된다.

1. 동공 영역 추출 및 중심 검출

1.1 동공 영역 추출

동공의 중심을 검출하기 전에 얻어진 눈 영상에 대해 동공의 영역만을 추출해 해당 영역만 처리를 한다면 훨씬 효과적일 것이다. 왜냐하면 영상 전체에 대해 컨볼루션 처리나 Hough 변환, 타원 fitting 등의 알고리즘을 적용하면 시스템에 불필요한 부하를 줄 가능성이 높고 시스템의 전체적인 속도에 큰 영향을 미칠 수 있기 때문이다. 따라서 본 시스템은 동공의 중심을 검출하기 전에 동공 영역을 먼저 추출하는 과정을 포함하고 있다.

처음 얻어진 눈 영상에 전체에 대해 이진화를 수행하면 영상에는 원하지 않는 눈썹이나 광원의 반사, 혹은 머리카락 등에 의한 노이즈가 남게 된다. 이러한 노이즈를 간단히 없애고 계산상의 복잡도를 낮추기 위해 저대역 통과 필터라 할 수 있는 가우시안 피라미드 분해 방식이 이용되었다. 2단계의 다운-샘플된 영상은 실제 입력 영상의 1/16 크기 영상으로, 여기에 추가로 미디안 필터링이 적용된다. 이 단계는 저대역 통과 필터 후에도 남아있는 노이즈를 제거하기 위한 것이다. 미디안 필터링 적용 후 남는 영역들은 머리카락이나 동공, 콧구멍에 의한 픽셀 값들이다. 이렇게 예상되는 영역들 중 몇 가지를 다음 그림4에서 표현하고 있다.

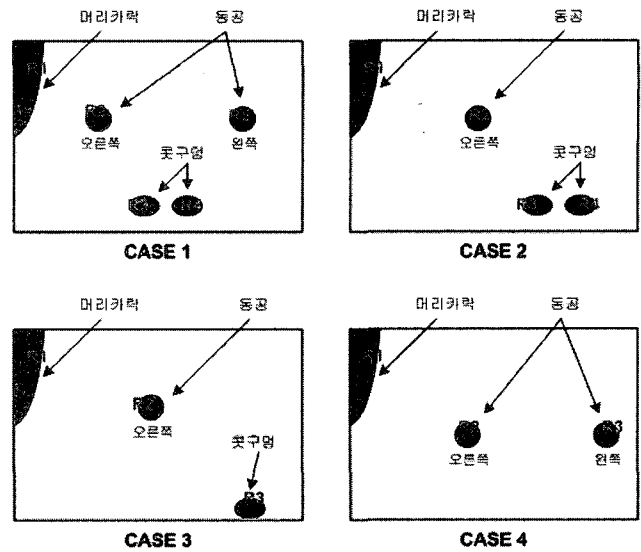


그림 4. 남아있는 영역 예시도
Fig. 4. Example Cases of Remaining Regions

위에 나타난 바와 같이 앞의 두 단계를 거친 눈 영상은 주로 CASE 3에서처럼 머리카락 영역 (R1), 동공 영역 (R2), 콧구멍 영역(R3) 등으로 나뉘게 된다. 이 영역들 중 동공 영역을 추출해 내기 위해 본 시스템은 프로젝션 히스토그램을 사용하였다. 프로젝션 히스토그램이란 특정한 방향으로의 픽셀 값의 누적 결과이다. 각 경우의 예시에서도 볼 수 있듯이 검출된 영역들은 x축 방향으로 서로 겹치지 않는다는 것을 알 수 있다. 따라서 먼저 x축 방향으로 프로젝션 히스토그램을 수행하여 각 영역을 나누고, 다음으로 y축 방향의 프로젝션 히스토그램을 수행하여 분류된 영역들 중 동공 영역을 판단한다. 이 때 구분된 영역들에 대해 각각 그 x축 범위에서만 y축 프로젝션 히스토그램을 수행하기 때문에 y축 방향으로 픽셀들이 겹치는 것은 무관하다. y축 프로젝션 히스토그램 결과에서 머리카락은 y축 방향의 누적 값이 동공과는 확연히 구별되므로 간단히 동공이 아니라고 판별할 수 있다. 또 콧구멍은 사용자가 카메라를 똑바로 쳐다보고 있다는 가정 하에 항상 동공보다 아래쪽에 위치하므로, y축 방향의 누적 픽셀 값들이 나타나는 순서에 의해 제거될 수 있다. 이러한 과정을 거쳐 찾아낸 동공 영역 추출 결과는 다음 그림5에 표시되어 있다.

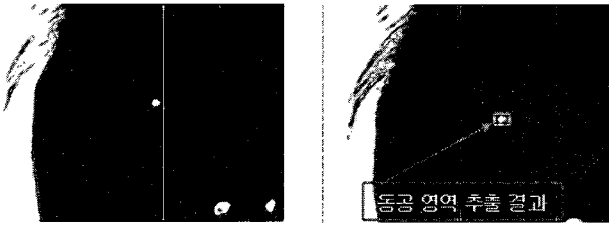


그림 5. 동공 영역 추출 결과
Fig. 5. Extraction Result of Pupil Region

1.2 동공 중심 검출

동공 영역을 추출한 후에는 효율적으로 영역 안에서만 동공의 중심을 찾는 단계를 수행할 수 있다. 여기에는 Hough 변환이나 active contour, coordinate averaging^[10] 등 원이나 타원을 찾는 알고리즘들이 수행될 수 있다. 본 시스템은 여러 알고리즘 중 효과적으로 타원의 방정식을 찾아내는 "Direct Least Squares Fitting of Ellipses"^[11] 방법을 사용하여 동공의 중심을 검출해 냈다.

다음 그림 6은 추출된 동공 영역 내에서 위의 알고리즘을 이용하여 최적화된 타원을 찾고 그 중심을 표시한 것이다. 이를 이용하면 주어진 데이터 집합에서 어떤 값 이상으로 떨어진 데이터는 무시하고 타원을 생성하므로, 동공이 눈썹 등에 가려서 완전히 보이지 않을 때에도 비교적 정확한 결과를 얻을 수 있는 장점이 있다.

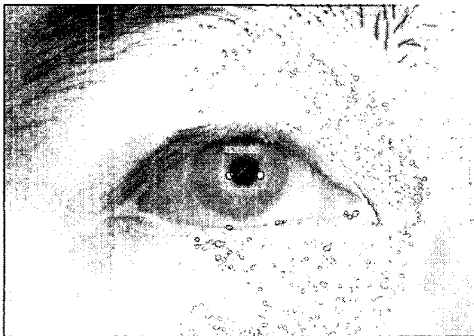


그림 6. 동공 중심 검출 결과
Fig. 6. Result of Pupil Center Detection

2. 반사점 중심 검출

반사점 중심 검출은 앞서의 동공 중심 검출 방법과 매우

유사하게 진행된다. 먼저 시스템은 전 단계에서 찾은 동공의 중심을 기준으로 100x50 픽셀에 해당하는 가상의 검색 영역을 생성한다. 모니터의 양 쪽에 부착되는 두 적외선 광원은 항상 동공의 주위에 반사점을 만들기 때문에 이러한 제한 내에서 영상 처리를 수행하는 것은 앞서의 동공 영역 추출 단계와 마찬가지로 시스템의 효율을 높일 수 있다. 하지만 동공과 달리 두 반사점은 정확한 중심을 찾아내기에는 상대적으로 크기가 작고, 동공의 경계는 반사점을 검출해 낼 때 방해가 된다. 따라서 본 시스템은 가우시안 피라미드 분해 방법을 사용해 검색 영역을 1단계 업-샘플 하여 정확한 반사점 중심 검출을 수행한다. 이 때 동공의 경계는 저대역 통과 필터가 반사점에 비해 낮은 픽셀 밝기 차이인 동공의 경계를 제거해 준다.

이 후의 영상에는 Canny 경계 검출이 적용된다. Canny 경계 검출의 Hysteretic counter 이진화 과정은 노이즈 성분의 경계는 여러 개의 경계 조각으로 나뉘어 검출되지 않고, 강한 경계 성분이나 서로 연결되어 있는 경계만을 검출한다. 따라서 Canny 경계 검출을 이용하면 두 반사점의 경계만이 정확히 검출된다. 이 때 반사점의 중심을 구하기 위해 다른 특별한 알고리즘을 사용하지 않고, 경계 검출 결과에 프로젝션 히스토그램을 x, y축으로 적용하여 각 중심만을 택하면 쉽게 두 반사점의 중심을 얻을 수 있다. 이 과정은 다음 그림 7에 설명되어 있다.

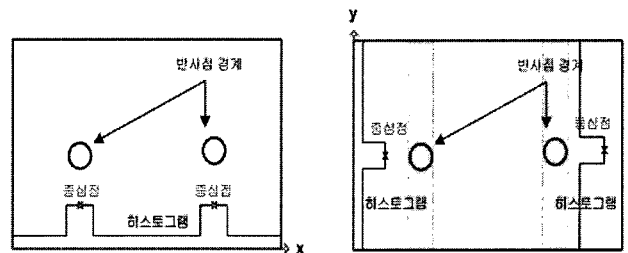


그림 7. 반사점의 프로젝션 히스토그램 (x축, y축)
Fig. 7. Projection Histogram of Glints (x-axis, y-axis)

아래 그림 8은 지금까지 설명된 방법을 사용해 실제 반사점 중심을 검출해 본 결과이다. 그림에서처럼 이 알고리즘을 사용하면 두 개의 반사점 경계만이 정확히 검출 된다는 것을 확인할 수 있다.

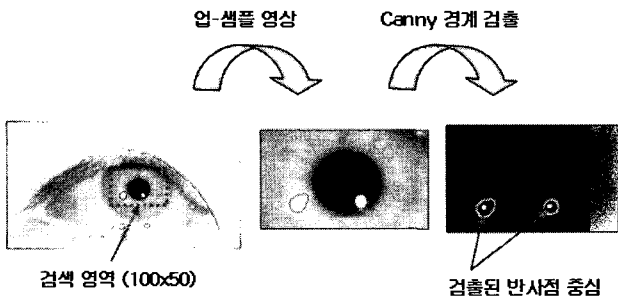


그림 8. 반사점 중심 검출 결과
Fig. 8. Result of Glint Center Detection

3. 검출 결과 검증

앞에서 얻은 동공과 반사점 중심의 정확도를 알아내기 위해 검증 과정을 거쳤다. 시선 추적 시스템에서는 작은 픽셀 단위의 오차라도 예측된 최종 시선 결과에 큰 영향을 미칠 수 있다. 우리는 얻어진 결과를 확인하기 위해 정해진 위치의 동공과 반사점을 포함한 푸르킨에 영상과 유사한 그림을 제작했다. 이 그림을 본 시스템에 입력한 결과, 그림 9에서 보여 지듯이 실제 값과 검출된 값 사이의 오차가 1 픽셀 이하로 신뢰할 수 있는 결과를 나타냈다.

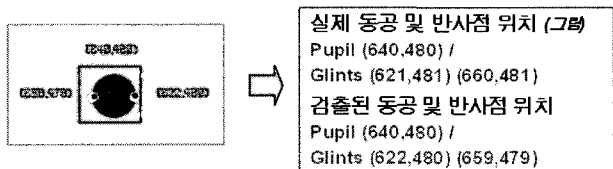


그림 9. 검출 결과 검증
Fig. 9. Confirmation of Detection

IV. 시선 추적 알고리즘

본 시스템에서 제안된 시선 추적 알고리즘은 간단하다. 시스템은 단지 입력된 눈 영상의 2차원 정보와 개인별로 캘리브레이션 된 데이터를 가지고 시선 추적을 하게 된다. 카메라나 적외선 광원의 재배치, 사용자의 머리 움직임은 입력된 영상이 푸르킨에 영상, 즉 동공과 두 개의 반사점을 포함한 영상이면 허용된다.

대부분의 시선 추적 장치들과 달리, 본 시선 추적 시스템은 복잡한 카메라 캘리브레이션 과정을 필요로 하지 않는다. 보통의 시선 추적 시스템은 사용자의 눈이나 동공, 각막 곡률 중심 등의 3차원 정보를 필요로 하기 때문에, 시스템을 사용하기 전에 카메라 좌표계와 광원의 좌표계, 그리고 스크린 좌표계 사이의 관계를 규정하는 캘리브레이션이 선행되어야 한다. 이 과정은 해당되는 시스템 자체에 특화되어 있고, 때때로 매우 복잡하여 수행하기 힘들거나 특별한 장치들을 요구한다. 하지만 본 시선 추적 시스템은 사용자의 불편을 초래하지 않는 간단한 사용자 캘리브레이션 과정을 제외하고는 어떠한 사전 수행도 필요치 않다. 따라서 본 시스템은 간편한 동시에 여타 시스템이 카메라나 광원의 물리적 움직임이 생겼을 경우 각각의 좌표계를 규정하는 행렬식이 변하기 때문에 처음부터 다시 캘리브레이션을 수행해야 하는 데 반해, 시스템이 반사점이 포함된 눈 영상을 얻을 수 있다면 상관없이 동작하게 된다. 또한 사용자의 머리가 움직일 때도 두 반사점의 위치는 항상 일정하므로 시스템에 영향을 주지 않는다. 다만 보다 정확한 결과를 얻기 위해서는 장비나 사용자의 머리가 움직였을 시 사용자 캘리브레이션을 다시 수행할 수 있다. 이는 다른 시스템에도 공통적인 사항이지만 본 시스템은 특히 간편한 사용자 캘리브레이션 과정을 지원함으로써 불편함을 최소화 하였다.

1. 시선 추적 알고리즘

본 시스템의 시선 추적 알고리즘은 입력된 눈 영상의 2차원 위치 정보만을 이용하기 때문에 간단한 2차원 방식

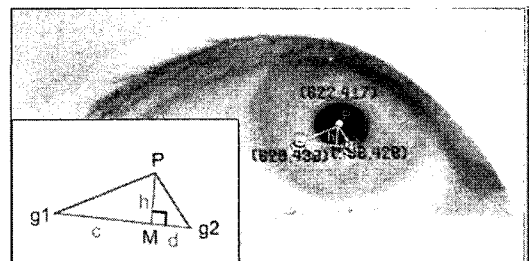


그림 10. 시선 추적 알고리즘의 특징점 관계
Fig. 10. Feature Relations of Gaze Tracking Algorithm

이라 부른다. 스크린 상에서 사용자가 바라보고 있는 주시점을 알아내기 위해서 아래 그림 10에 표시된 것과 같이 동공의 중심 P와 두 반사점 중심 g1, g2로 이루어진 세 개의 특징점 사이의 거리 관계를 이용한다.

두 반사점 g1과 g2를 잇는 선에 동공의 중심 P로부터 수선의 발을 내린 점을 M이라고 한다. h는 P와 M을 잇는 선의 길이이고 c와 d는 각각 g1과 M, M과 g2를 잇는 선의 길이를 나타낸다. 이 h, c, d의 길이와 시선간의 관계가 선형이라고 가정하면, h는 사용자가 바라보고 있는 스크린 상 y축의 주시점 위치를 예측하는 데 사용될 수 있다. 마찬가지로 c나 d 값은 주시점의 스크린 상 x축 위치를 예측할 수 있다. 둘 중 어느 값을 사용하여도 상관없지만 본 논문에서는 주로 c값을 사용하기로 한다. 다음에 설명될 사용자 캘리브레이션 과정을 통해 각 개인별 초기 값을 가지고 시스템은 사용자의 스크린 상 주시점을 판별하게 된다.

2. 사용자 캘리브레이션

시스템을 사용하기에 앞서, 각 사용자는 사용자 캘리브레이션 과정을 수행해야 한다. 이 과정은 모든 시선 추적 시스템이 포함하는 과정으로 본 시스템에서는 이것을 보다 간편하게 구성하였다. 이를 통해 시스템은 사용자 눈의 위치나 회전 정도 등의 사용자 의존적인 요소들을 초기화시킬 수 있고, 이로 인해 앞서 세 특징점들 사이의 관계를 재정의할 수 있다. 대부분의 시선 추적 시스템은 시스템 사용 전에 수 개 혹은 수십 개의 물체를 여러 단계에 걸쳐 쳐다봐야 하는 반면, 본 시스템은 단 두 단계로만 이루어져 시간을 절약하고 사용자의 편의를 높일 수 있다. 이와 유사한 두 단계의 사용자 캘리브레이션은 먼저 T. Ohno의 시선 추적 시스템 [8]에 의해 제안되었다. 이 시스템은 사용자의 실제 바라본 점과 시스템이 예측한 점 사이의 오차를 두 가지 파라미터로 저장하여 활용하는데, 본 시스템은 각 특징점들 사이에 관계된 초기 값을 규정하기 위해 사용한다. 따라서 사용자 캘리브레이션 후 사용자는 자신에게 적용되는 시스템의 해상도를 즉각 알 수 있다.

사용자는 첫 단계에서 좌측 상단에 위치한 점을 바라보도록 유도되고, 다음으로 우측 하단에 위치한 점을 바라보

게 된다. 이 값을 통해 각 특징점들 사이의 관계가 정의된다. 이 과정은 다음의 그림 11에 표시되어 있다.

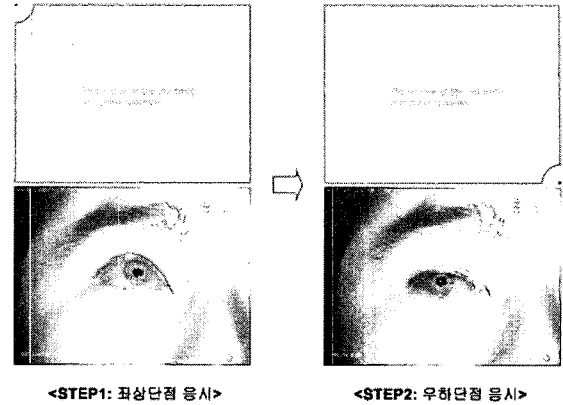


그림 11. 사용자 캘리브레이션
Fig. 11. Personal Calibration

3. 시스템 해상도

사용자 캘리브레이션 결과로 c와 h의 초기 값이 각각 $init_c1$, $init_c2$, $init_h1$, $init_h2$ 로 정해지면, 전체 시스템의 해상도는 (1)식에 의해 간단히 계산될 수 있다. 이 때 해상도는 사용자가 스크린을 바라보았을 때 예측되는 주시점의 해상도이고, 사용자의 위치나 안구의 움직임 정도에 따라 서로 다르게 나타난다. 초기 값의 뒤에 붙어 있는 숫자는 사용자가 스크린의 두 코너 점을 바라보았을 때의 최소 값과 최대 값을 차례로 나타내고 width와 height는 모니터의 해상도를 나타낸다.

$$\begin{aligned} resolution_x &= \frac{width}{|init_c1 - init_c2|} \\ resolution_y &= \frac{height}{|init_h1 - init_h2|} \end{aligned} \quad (1)$$

사용자마다 다르지만 보통 사용자 캘리브레이션을 수행하면 c와 h의 최대 및 최소 초기 값 차이는 20~30픽셀 정도로 결정된다. 따라서 시스템 해상도는 XGA 스크린 (1024x768)의 경우 x축으로 34.13픽셀부터 51.2픽셀까지

범위를 가질 수 있고, y축으로는 25.6픽셀부터 38.4픽셀의 범위를 가질 수 있다. 해상도는 고화질의 카메라나 줌 렌즈 사용, 혹은 스크린 해상도를 낮추면 개선될 수 있지만, 고화질 카메라 사용은 연산 시간을 늘릴 수 있고 줌 렌즈는 카메라 시야각을 줄여서 사용자의 머리 움직임을 제한할 수 있다.

V. 시스템 구성

본 시스템은 Point Grey사의 Dragonfly™ 카메라를 사용하고 광원으로는 두 개의 적외선 LED 집합을 사용한다. 카메라는 XGA (1024x768)급의 해상도를 15fps로 지원하고 시스템 구동 시에 10~12fps 정도의 성능을 보여주어 실시간으로 동작한다. 적외선 LED의 총 전력은 사용자의 안전을 위해 2와트 이하로 제한하였고 파장은 880nm대의 근적외선을 사용한다. PC는 PENTIUM 4 3.2GHz, 1GB의 DDR RAM과 17인치 TFT-LCD 모니터를 사용한다.

응용 프로그램으로 Microsoft Visual C++ 6.0 외에 Dragonfly™ SDK, OpenCV 3.1b 버전을 사용하여 개발되었다. 전체 시스템의 구조도와 실제 사진은 다음의 그림 12와 13에 나타나 있다.

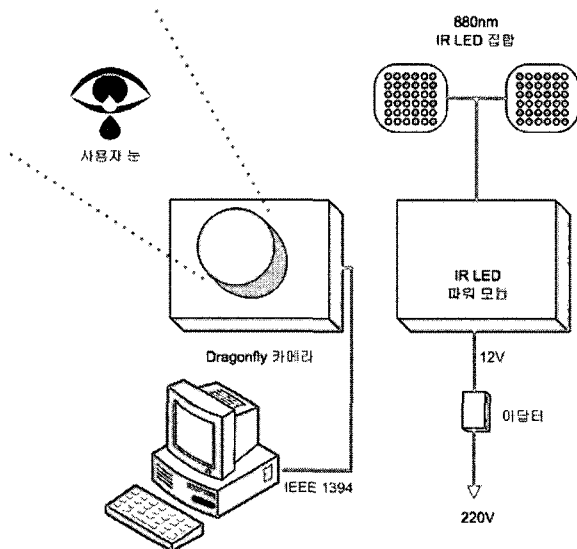


그림 12. 시스템 구조도
Fig. 12. System Configuration

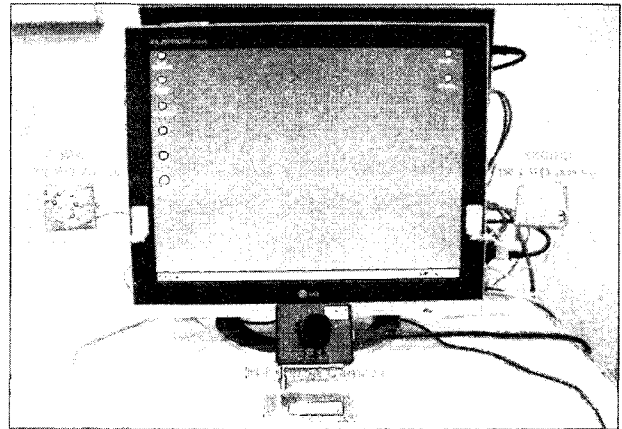


그림 13. 시스템 외관
Fig. 13. System Appearance

VI. 실험 결과

본 시스템의 정확도는 스크린 상 위치를 알고 있는 9개의 고정점을 사용자가 바라보게 함으로써 측정되었다. 3명의 피실험자로 하여금 이 9개의 점을 차례대로 바라보게 하였고, 전체를 5번 반복하여 그 결과를 기록하였다. 그림 14는 주시점의 x좌표 예측 결과와 실제 위치와의 차이가 픽셀 단위로 기록되었고, 그림 15는 주시점의 y좌표 예측 결과와 실제 위치의 차이가 기록되었다. 모든 결과는 5번 반복된 실험의 평균치를 나타낸다. 대부분의 오차 값은 2

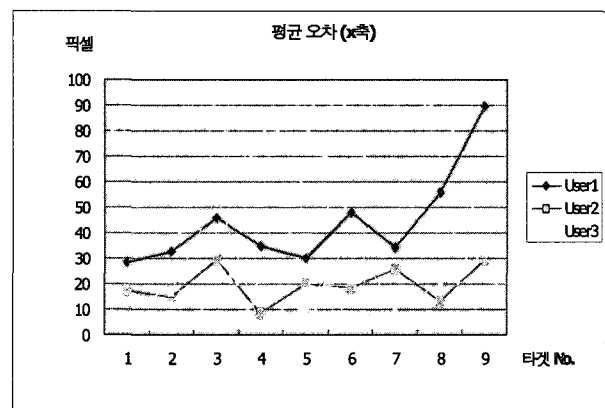


그림 14. 평균 오차 (x축)
Fig. 14. Average Error (x-axis)

0~30픽셀에 몰려 있는 것을 알 수 있고, 이는 곧 시스템의 평균 해상도와도 일치한다.

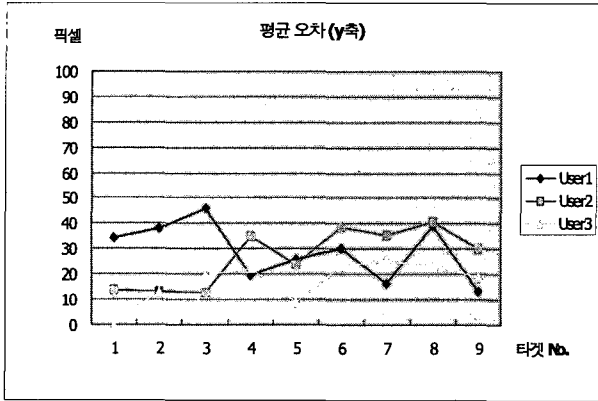


그림 15. 평균 오차 (y축)
Fig. 15. Average Error (y-axis)

각 사용자별 표준 편차는 그림 16에 표시되어 있다. 결과를 보면 사용자 1의 결과가 다른 사용자에게 비해 다소 편차가 심한 것으로 보이지만, 모든 결과는 20 이하의 비교적 안정적인 결과를 보이고 있다.

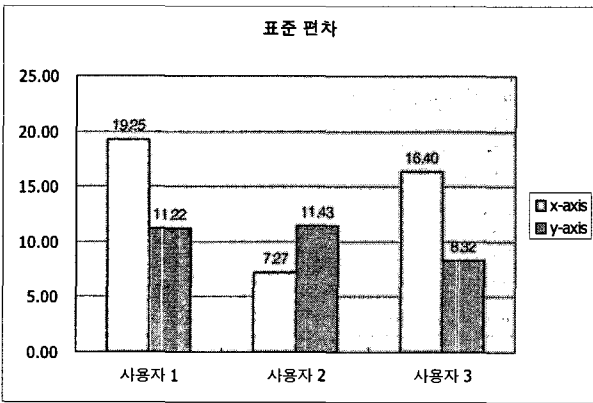


그림 16. 각 사용자별 표준 편차
Fig. 16. Standard Deviation of Each User

시스템의 평균 오차는 다음 그림 17의 방법으로 측정되었다. 그림 17에서 사용자는 스크린으로부터 600mm 떨어진 것으로 가정되었다. 실제 스크린 상의 주시점과 예측된 주시점 사이의 각도는 그림 17에 포함된 식으로 알 수 있

고, 평균 오차는 x축으로 0.9도, y축으로 0.82도를 기록해 비교적 정확한 결과를 보였다.

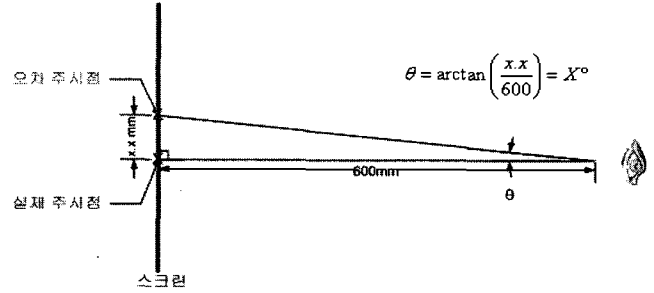


그림 17. Total Average Error Analysis
Fig. 17. 총 평균 오차 분석

VII. 응용 분야

본 논문에서 제안한 시선 추적 시스템을 이용해 개발된 응용 프로그램은 다음과 같다.

1. 날짜 선택 프로그램

날짜 선택 프로그램은 전체 스크린에 4x4 격자를 그리고 각 칸에 숫자를 배치하여 날짜를 타이핑 할 수 있는, 키패드 대체 용 프로그램이다. 날짜의 선택은 숫자를 1초간 바라보는 것으로 구현되어 있고, 눈 깜박임 등 다른 동작으로 대체될 수 있다.



그림 18. 날짜 선택 프로그램
Fig. 18. Date Selection

2. 시선 기반 마우스

시선 기반 마우스는 마우스 후킹을 이용해 전체 시스템의 마우스를 시선으로 대체하여 기존의 마우스가 가지는 각종 기능을 시선으로 대신하는 프로그램이다. 이를 응용하면 장애인을 위한 인터페이스로 뿐만 아니라, 일반인들의 관심도를 분석하여 적절한 피드백을 주는 AUI (Attentive User Interface)로써 활용할 수 있다.

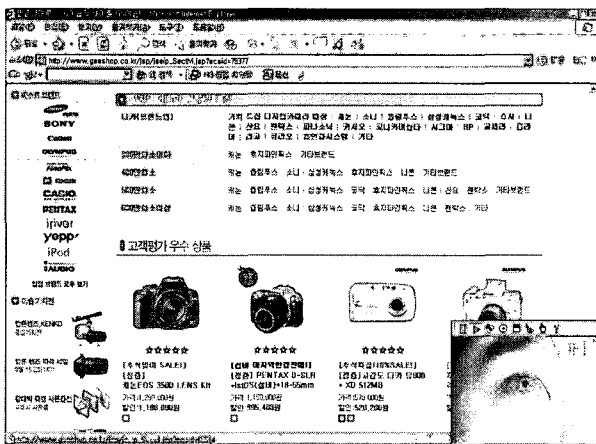


그림 19. 시선 기반 마우스
Fig. 19. Gaze-Based Mouse

Ⅷ. 결 론

본 논문에서는 간편한 2차원 시선 추적 시스템이 제시되었다. 이 시스템의 장점은 다음과 같다. 첫째, 한 대의 카메라와 두 개의 적외선 광원만이 사용되어 많은 자원을 필요로 하지 않고, 시스템 시작 시 처음부터 고정되거나 동작 중 움직이지 말아야 하는 등 제한점이 줄어들었다. 시스템은 모든 처리 과정을 거쳐 초당 10~12프레임의 실시간으로 동작한다. 둘째, 매우 간편한 사용자 캘리브레이션 과정을 포함한다. 이는 시스템이 머리 움직임이나 다른 조건에 의해 안 좋은 결과를 나타낼 때, 사용 중 다시 캘리브레이션을 하는 것에 부담을 느끼지 않을 수 있다는 큰 장점이다. 또한 시스템이 동작 중에 카메라나 적외선 광원을 더 좋은

영상을 얻기 위해 움직일 수 있으며, 사용자의 머리 움직임도 눈 영상을 획득할 수 있는 한 허용되기 때문에 매우 유연한 시스템이라고 할 수 있다.

참 고 문 헌

- [1] S. W. Shih and J. Liu, "A Novel Approach to 3-D Gaze Tracking Using Stereo Cameras," IEEE Trans. On Systems, Man, and Cybernetics Part B: Cybernetics, vol. 34, no. 1, Feb. 2004.
- [2] J. Gips, P. Olivieri, and J. Tecce, "Direct Control of the Computer through Electrodes Placed Around the Eyes," Proc. Fifth Int. Conf. Human-Computer Interaction, pp. 630-635, 1993.
- [3] Y. Ebisawa, "Realtime 3D Position Detection of Human Pupil," IEEE Int. Conf. On Virtual Environments, Human-Computer Interfaces, and Measurement Systems, July 2004.
- [4] T. E. Hutchinson, J. K. P. White, W.M. Martin, K. C. Reichert, and L. A. Frey, "Human-Computer Interaction Using Eye-Gaze Input," IEEE Trans. On Systems, Man, and Cybernetics, vol. 19, pp. 1527-1534, Nov./Dec. 1989.
- [5] T. Cornweert and H. Crane, "Accurate Two-Dimensional Eye Tracker Using First and Fourth Purkinje Images," J. Opt. Soc. Amer., vol. 63, no. 8, pp.921-928, 1973.
- [6] L. Bour, "DMI-Search Scleral Coil," Dept. Neurology, Clinical Neuro-Physiol., Academic Medical Cent., AZUA, Amsterdam, Netherlands, Tech. Rep. H2-214, 1997.
- [7] <http://www.seeingmachines.com/face/ab.htm>
- [8] T. Ohno, N. Mukawa, A. Yoshikawa, "FreeGaze: A Gaze Tracking System for Everyday Gaze Interaction," ETRA02, 2002.
- [9] D. Beymer and M. Flickner, "Eye Gaze Tracking Using an Active Stereo Head," Proc. Of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003.
- [10] A. W. Fitzgibbon, M. Pilu, R. B. Fisher, "Direct Least Squares Fitting of Ellipses," Proc. of ICPR96, 1996.
- [11] Chaudhuri, B. B., Kundu, P., "Optimum Circular Fit to Weighted Data in Multidimensional Space," Pattern Recognition Letters, 14, 1-6, 1993.

 저 자 소 개

**박진우**

- 2004년 : 홍익대학교 전자전기공학 학사 졸업
- 2006년 : 연세대학교 전기전자공학 석사 졸업 (한국과학기술연구원 학연과정)
- 2006년 ~ 현재 IBM 유비쿼터스 컴퓨팅 연구소 연구원
- 주관심분야 : HCI, 혼합현실, 유비쿼터스컴퓨팅

**권용무**

- 1980년 : 한양대학교 전자공학과 학사 졸업
- 1983년 : 한양대학교 전자공학과 석사 졸업
- 1992년 : 한양대학교 전자공학과 박사 졸업
- 1983년 ~ 현재 : KIST 영상미디어연구센터 연구원, 선임연구원, 책임연구원
- 주관심분야 : 실감미디어, 3차원 영상, 인터랙션미디어

**손광훈**

- 1983년 연세대학교 전자공학과
- 1985년 Univ. of Minnesota, Dept. of EE
- 1991년 North Carolina State Univ. Dept. of ECE
- 2005년 연세대학교 전기전자공학 교수
- 주관심분야 : 3차원 동영상 CODEC, 3차원 혼합현실 콘텐츠 기술, 3D 얼굴인식 기술